基于 SEF-GHEI 及协同表示的步态识别研究 *

徐守坤 [1†]，邱　亮 [1]，石　林 [1]，李　宁 [1, 2]

(1. 常州大学 信息科学与工程学院 数理学院，江苏 常州 213164; 2. 福建省信息处理与智能控制重点实验室(闽江学院)，福州 350108)

摘　要：将协同表示方法应用于步态识别中可以解决稀疏表示方法计算耗时的问题，但提取步态特征采用的 GEI 算法没有考虑步态内部轮廓边界信息，导致识别率不高。针对此问题，提出使用融合 Hog 和 GEI 算法的方法提取步态特征，在此基础上使用协同表示的方法训练，再通过计算测试样本的最小重构误差进行分类。实验结果表明，该方法在单一视角下步态识别准确率平均提高了 1.315%，以及跨视角下步态识别准确率平均提高了 6.51%，说明本方法是可行的。

关键词：稀疏表示；协同表示；GEI 算法；Hog 算法

中图分类号：TP391.41　　doi: 10.3969/j.issn.1001-3695.2018.01.0052

Gait recognition based on SEF-GHEI and synergy representation

Xu Shoukun[1†], Qiu Liang[1], Shi Lin[1], Li Ning[1, 2]

(1. *School of Mathematics & Physics*, *School of Information Science & Engineering*, *Changzhou University*, *Changzhou Jiangsu* 213164, *China*; 2. *Fujian Provincial Key Laboratory of Information Processing & Intelligent Control* (*Minjiang College*), *Fuzhou* 350108, *China*)

Abstract: Applying collaborative representation method to gait recognition research can solve the problem of time-consuming calculation of sparse representation. While the GEI algorithm for extracting gait features does not consider gait internal outline information, resulting a low recognition rate. To address this problem, this paper proposed a novel method to extract gait features by using a fusion Hog and GEI algorithm. On the basis of this method, train the samples by using collaborative representation method and classify the samples according to the minimum reconstruction error. The experimental results show that the accuracy of the proposed method is 1.315% higher than that of the collaborative representation in a single view, and the accuracy of gait recognition is increased by 6.51% over the cross view.

Key words: sparse representation; collaborative representation; GEI algorithm; Hog algorithm

0　引言

步态识别是根据人行走的频率、相位以及胳膊摆动等来进行识别，与其他识别技术相比（人脸识别、语音识别、虹膜识别等），由于步态识别具有远距离识别、难以模仿以及易采集等特点，使其成为一种无可代替的生物识别技术，并且在安全敏感环境（如机场、银行、商场、火车站等）中具有广阔的发展和应用前景[1~5]。

目前步态识别主要分为两大类，一类是基于模型(model-based)的方法，一类是基于非模型(model-free)的方法[6~9]。基于模型的方法主要是通过对建立的人体运动模型进行跟踪分析获取参数，利用获取的参数作为步态特征进行匹配。主要有 Lee 等人[10]提出使用椭圆模型的方法、Bobick 等人[11]提出基于静态步态参数的表示方法。但上述两个方法均对步态视屏序列图像质量要求高，并且在建立模型时计算量大，难以实现较理想的结果。基于非模型的方法不需要预先假定模型，直接提取图像序列中运动目标轮廓产生的时空特性作为步态特征。主要有李占利等人[12]提出基于步态高斯及稀疏表示的步态识别，识别效果显著并在跨视角下具有一定的鲁棒性，但是在计算分类的过程中存在着耗时的问题，在较复杂的场景中满足不了实时性的要求。杨旗等人[13]提出基于稀疏表示的步态识别，识别准确性高且速度快。但步态特征提取选用的是主动能量图，该能量图只考虑了步态的动态信息而步态的静态信息被完全的丢弃。李占利等人[14]提出基于协同表示的步态识别，识别效果良好且计算速度快。但步态特征提取采用的是 GEI 能量图，由于 GEI 能量图只能捕获轮廓外部边界信息，忽略了轮廓内部边界信息的

收稿日期：2018-01-05；**修回日期：**2018-02-11　　**基金项目：**闽江学院福建省信息处理与智能控制重点实验室开放课题资助项目（MJUKF201740）

作者简介：徐守坤（1972-），男（通信作者），吉林蛟河人，教授，博士，主要研究方向为人工智能、普适计算等（16106406@smail.cczu.edu.cn）；邱亮（1993-），男，硕士研究生，主要研究方向为数据处理、图像识别；石林（1979-），男，硕士，副教授，主要研究方向为数据处理、图像识别；李宁（1974-），男，博士，副教授，主要研究方向为数据与信息处理.

原因，对步态的识别率有着一定的影响。众所周知，梯度方向直方图(Hog)不仅能够很好地描述轮廓图像的内外边界信息，还能对图像轮廓重合的边界信息进行提取，比 GEI 算法提取的轮廓边界信息更加详细，因此对跨视角下的步态识别效果更好[15~19]。

基于上述原因本文提出采用 Hog 和 GEI 算法融合的方法，即 SEF-GHEI 能量图作为步态特征，采用基于协同表示的分类器进行识别。

1　基于 SEF-GHEI 及协同表示的步态识别方法原理与基本过程

SEF-GHEI 能量图是由一个完整周期内步态图像的 Hog 特征相加求平均计算得到的，该能量图不仅反映出步态周期内轮廓图像的变化情况以及步态边缘变化的特征，且减少了步态数据量。在对提取到的 SEF-GHEI 能量图进行分类时需解决如下两个问题：a）由于该能量图维数过高的原因，直接分类会遇到计算耗时、运算复杂等问题，所以本文采取 PCA 方法进行降维；b）由于不同对象之间的 SEF-GHEI 能量图具有一定的相似性以及每个对象的 SEF-GHEI 能量图数量少的原因，使用该类对象的训练样本来表示测试样本会产生误差，所以本文以 SEF-GHEI 能量图为基本的步态特征数据，通过使用协同表示分类器进行分类识别。

1.1　基本原理

1.1.1 SEF-GHEI 能量图

人在行走过程中胳膊的摆动幅度、走路的频率、步幅等特性是存在周期性变化。因此，为了能够更好地描述人的步态特点，本文从步态周期序列图像中提取每个人的步态特征。SEF-GHEI 能量图像是一步态周期内每帧步态剪影图像 Hog 特征的平均，该幅图像反映了轮廓的主要形状、内外部边界信息以及一个步态周期内形状的变化，不仅减少了步态数据量，而且对每帧图像中的噪声也并不敏感。SEF-GHEI 能量图定义如下：

$$g(x,y,b)=\frac{1}{N}\sum_{t=1}^{N}h_t(x,y,b) \qquad (1)$$

其中：N 表示步态周期内的步态剪影图像序列帧数；$h_t(x,y,b)$ 表示第 t 帧步态剪影图像坐标 (x,y) 像素值 b 方向的 Hog 特征，其中 $b\in\{1\cdots 36\}$ 表示梯度方向直方图 bin 的索引号。图 1 为同一对象不同状态下的 SEF-GHEI 能量图。

1.1.2 协同表示

基于稀疏表示的分类是假定每类的测试样本可以用训练样本线性组合，且每一类的样本足够多，即构建的字典是过完备字典[20]。但是在实际应用中，每一类的训练样本数目较少，因此构建的字典是不完备的。Zhang 等人[21]提出通过使用协同表示分类方法(CRC)将每类训练样本构建成一个字典 A 来测试样本 y 的分类，很好地解决了这一问题，并提出正则化最小二乘分类方法。该方法与 SRC 模型：

$$(\hat{x})=\arg\min_x\|x\|_1 \quad s.t.\|y-Ax\|_2<\varepsilon \qquad (2)$$

相比，CRC 方法没有使用约束条件，并且将 l_1 范数替换成 l_2 范数，使分类问题变成了最小二乘问题。为了降低求解的计算成本，本文使用正则化最小二乘方法，即

$$(\hat{x})=\arg\min_x\left\{\|y-Ax\|_2^2+\lambda\|x\|_2^2\right\} \qquad (3)$$

其中：λ 表示为正则化参数，通过调节其参数可以使最小二乘解更稳定，以及让 x 达到一定的稀疏性。

从式(3)中可以推导出 $\hat{x}$ 的解：

$$(\hat{x})=(A^TA+\lambda\cdot A)^{-1}A^Ty \qquad (4)$$

令 $P=(A^TA+\lambda\cdot I)^{-1}A^T$，则 P 与 y 不相关，可以先将 P 作为投影矩阵计算出来。在测试分类的过程中，通过 $\hat{x}$ 和 A 对测试样本 y 进行重构，并使用 $\|y-A_ix_i\|_2$ 的重构误差以及 l_2 范数 $\|x_i\|_2$ 的解来进行分类。

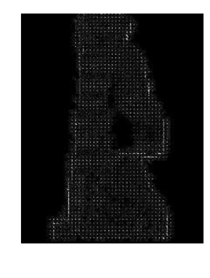
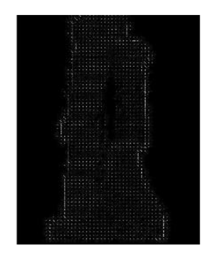
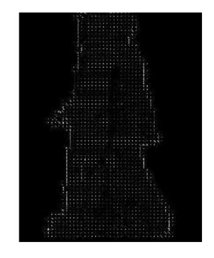




图 1　同一对象不同状态下的 SEF-GHEI 能量图

1.2　基本过程

本文首先使用 GEI 和 Hog 算法融合的方法对步态序列图像进行特征提取，即 SEF-GHEI 步态能量图，其次以 SEF-GHEI 步态能量图为基本输入数据，并进行训练得到投影矩阵 P，根据计算测试样本在矩阵 P 上的投影即正规化残差进行分类识别，得到识别结果。其整个步态识别流程如图 2 所示。

2 基于 SEF-GHEI 及协同表示的步态识别计算过程

2.1 获取 SEF-GHEI 能量图过程

a）步态轮廓提取。由于自适应混合高斯模型和帧间梯度信息的人体目标分割算法可以消除提取的目标图像的影子以及光照影响，所以本文采用该算法从步态视屏序列中提取人体目标图像并二值化，然后使用形态学等算子消除提取的轮廓图像中含有的噪声以及小空洞。

b）图像标准化。将二值化的步态轮廓图像以水平中心对齐的方式并按照一定的比例进行缩放，得到大小为 400×500 的中心化及归一化图像。

c）步态剪影提取。从标准化的步态轮廓图像中提取步态轮廓的剪影。

d）Hog 特征提取。将步态剪影图像划分为 8×8 像素的 cell，且 cell 之间不重叠，并统计 cell 不同梯度方向的 bin 数，将 cell 的特征描述符串联起来就可以得到该帧步态剪影的 Hog 特征。

e）SEF-GHEI 能量图获取。将完整步态周期内的图像根据式(1)计算得到 SEF-GHEI 能量图。

2.2 步态训练的计算过程

a）将所有训练样本的 SEF-GHEI 能量图以列的方式展开成列向量，然后把展开成列向量的训练样本构建成字典 $A=[A_1,A_2,\cdots,A_n]$，其中 A_i 表示第 i 类训练样本组成的列向量特征。

b）将矩阵 A 的列进行规范化计算，并得到 l_2 范数。

c)由于 SEF-GHEI 步态能量图的像素是 400×500，构建的字典 A 的维度为 20 000 维，维度较高且计算复杂的原因，本文使用 PCA 方法先对字典 A 进行降维处理。

d）通过公式 $(A^TA+\lambda I)^{-1}A^T$ 计算投影矩阵 P，即

$$P=(A^TA+\lambda\cdot I)^{-1}A^T \tag{5}$$

2.3 步态识别的计算过程

a）将所有测试样本的 SEF-GHEI 能量图以列的方式展开成列向量，并进行列规范化。

b）使用 y 表示采用 PCA 方法进行降维后的测试样本。

c）计算降维后的测试样本 y 在矩阵 P 上的投影值 x，即

$$x=Py \tag{6}$$

d）通过 $\frac{\|y-A_ix_i\|_2}{\|x_i\|_2}$ 计算正规化残差 r_i，即

$$r_i=\frac{\|y-A_ix_i\|_2}{\|x_i\|_2} \tag{7}$$

e）将 y 归类于残差最小的一个类：$identity(y)=\arg\min_i(r_i)$ 并输出识别结果。

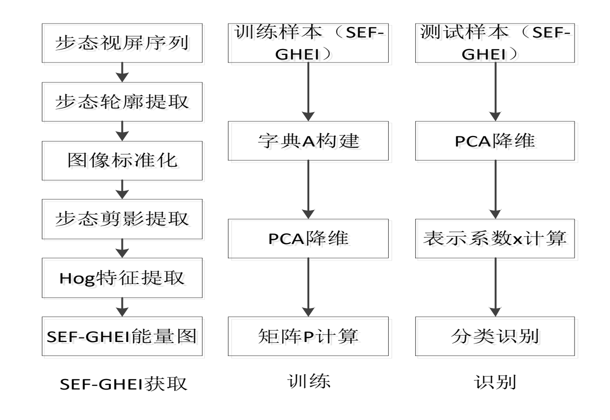


图 2　SEF-GHEI 获取、训练和识别的流程

3 实验结果分析

本文在 CPU 3.0 GHz，内存 8.0 GB，MATLAB R2010b 环境下进行实验的，实验数据使用的是 CASIA 数据库 DataSetB 包中的步态数据进行实验以及验证本文方法的识别性能，该包中有 124 个对象在 11 种视角下的背包、穿大衣以及正常的 3 种行走状态步态序列图像。本文选用该 124 个对象在 72°、90°以及 108°视角下的步态图像序列进行实验，且在实验过程中所使用的 PCA 降维方法本文将其累积贡献率均设置为 0.95。

3.1 正则化参数 λ 对步态识别率的影响

引入正则化参数 λ 是为了能够让表示系数 x 更稳定以及具有一定的稀疏性，所以本文在 λ 取值不同的情况下分别对单一视角以及跨视角下的背包、穿大衣和正常行走的 3 种状态的步态数据集进行了识别验证，如图 3~5 所示。

从图中可以看出，当正则化参数 $\lambda=0$ 时，在 72°视角、90°视角以及 108°视角下的三种步态数据的识别率都非常的低；当正则化参数 λ 的取值区间在 0.000 001~1 间时，在 72°视角、90°视角以及 108°视角下的三种步态数据的识别率都非常理想；当正则化参数 λ 的取值超过 1 时，在 72°视角、90°视角以及 108°视角下的三种步态数据的识别率都开始降低。所以本文在实验过程中，将正则化参数 λ 的取值定位 0.001。

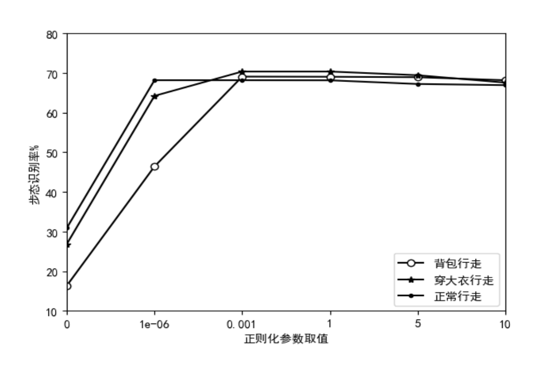


图 3　72°视角下正则化参数取不同值的识别率

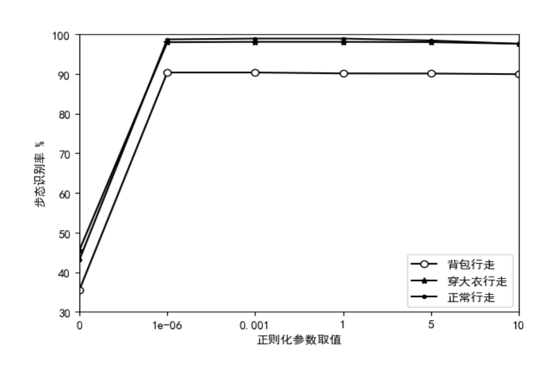


图 4　90°视角下正则化参数取不同值的识别率

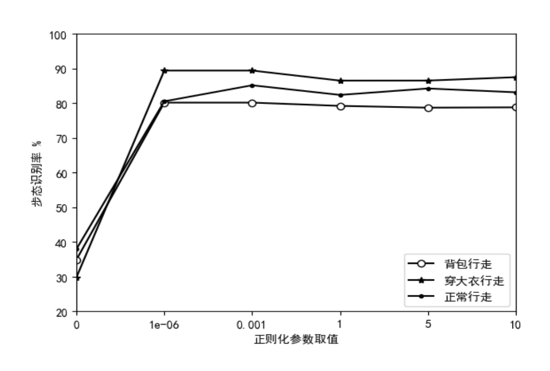


图 5　108°视角下正则化参数取不同值的识别率

3.2　单一视角下识别性能比较

每个对象在每种视角下有 2 个携带背包、2 个穿大衣以及 6 个正常行走的步态图像序列。本文分别从背包行走和穿大衣行走的步态图像序列中选取 1 个图像序列作为训练样本，剩下的图像序列作为测试样本；从正常行走的步态图像序列中选取 3 个图像序列作为训练样本，剩下的 3 个图像序列作为测试样本。由于 90°视角下步态图像序列的步态特征明显且易提取，所以本文使用该视角下的步态图像序列进行实验，并分别与基于步态高斯及稀疏表示分类方法和基于 GEI 能量图的协同表示方法进行比较，其中的识别时间是平均识别一个测试样本所用的时间，比较结果如表 1、2 所示。

表 1　90°视角下 SRC、NN 及本文算法的识别率

行走状态	算法及识别率 / %		
	GGI+SRC	GEI+CRC	本文算法
正常	98.26	98.93	98.94
穿大衣	94.76	97.18	98.11
携带背包	89.92	87.91	90.38

从表 1 中可以看出，本文算法在 90°视角下的三种行走状态的识别率均高于基于步态高斯图及稀疏表示的步态识别和基于 GEI 能量图协同表示的步态识别。本文算法与基于步态高斯及稀疏表示的步态识别准确率相比，本文方法使步态识别准确率平均提高了 1.5%左右，比基于协同表示的步态识别准确率平均提高了 1.13%。本文方法的识别率比基于步态高斯及稀疏表示的步态识别以及基于 GEI 能量图协同表示步态识别的识别准确率高的主要原因是提取步态能量图的方法不一样，在协同表示方法中采用的是 GEI 能量图来表述一个周期内的步态特征，而步态高斯能量图重在描述不同对象之间步态的差异性，由于 GEI 能量图和步态高斯能量图只考虑了人体轮廓的外部边界信息，没有考虑内部边界信息，对步态的识别率有着一定的影响。而本文提取步态特征采用的是 GEI 和 Hog 算法融合的方法，即 SEF-GHEI 能量图，该方法考虑了人体轮廓的内外部边界信息。因此，本文算法的识别率均高于基于协同表示的步态识别以及基于步态高斯图及稀疏表示的步态识别的识别率。

表 2　90°视角下 SRC、NN 及本文算法的识别时间

行走状态	算法及时间 / s		
	GGI+SRC	GEI+CRC	本文算法
正常	0.2075	0.0018	0.015
穿大衣	0.1222	0.0013	0.025
携带背包	0.1221	0.0013	0.0125

从表 2 中可以看出，本文采用的算法和基于 GEI 能量图的协同表示步态识别算法的识别时间均小于基于步态高斯及稀疏表示算法的识别时间，识别速度提高了 13.8~115.2 倍，其主要原因是式（4）中独立的部分即式（5）能提前计算出来，并且能够快速计算出测试样本的重构系数，而基于步态高斯图及稀疏表示的方法对于每次输入的测试样本都要计算一次 l_1 范数比较耗时。本文方法的识别时间高于基于协同表示步态识别的识别时间的主要原因是按式（1）计算并提取步态特征时，要先计算每张步态剪影的 Hog 特征，而 Hog 特征的计算以及提取比较耗时，但是本文方法的识别时间不超过 0.035 s，能满足实时性的要求。

3.3 跨视角下识别性能比较

为了验证本文算法在实际环境中的有效性，本文以 90° 视角下的三种行走状态步态序列图像作为训练样本，分别把 72° 和 108° 视角下的正常、穿大衣以及背包三种行走状态步态序列图像作为测试样本进行实验。实验结果如表 3~6 所示。

表 3 72° 视角下 SRC、NN 及本文算法的识别率

行走状态	算法及识别率 / %		
	GGI+SRC	GEI+CRC	本文算法
正常	66.13	62.37	68.18
穿大衣	69.76	70.58	71.30
携带背包	64.52	69.36	75.45

表 4 72° 视角下 SRC、NN 及本文算法的识别时间

行走状态	算法及时间 / s		
	GGI+SRC	GEI+CRC	本文算法
正常	0.485	0.0017	0.0294
穿大衣	0.1492	0.0013	0.0317
携带背包	0.1487	0.0013	0.0219

表 5 108° 视角下 SRC、NN 及本文算法的识别率

行走状态	算法及识别率 / %		
	GGI+SRC	GEI+CRC	本文算法
正常	82.93	71.51	85.18
穿大衣	88.31	80.24	89.42
携带背包	79.03	64.11	83.96

表 6 108° 视角下 SRC、NN 及本文算法的识别时间

行走状态	算法及时间 / s		
	GGI+SRC	GEI+CRC	本文算法
正常	0.3895	0.0017	0.0311
穿大衣	0.1502	0.0013	0.0253
携带背包	0.148	0.0013	0.0161

从表 3 和 5 中可以看出，本文算法在跨视角下的 72° 视角和 108° 视角的三种行走状态图像序列的识别率均高于另外两种算法，特别是在正常行走状态下的识别准确率平均提高了 9.74%，以及携带背包行走状态下的识别准确率平均提高了 12.97%。从表 4 和 6 可以看出，本文算法的识别时间相对于协同表示算法比较耗时，但是能满足实时性要求，比基于步态高斯图及稀疏表示算法的识别速度快 10 倍左右，并且在单一视角下和跨视角下，本文算法的识别率均高于基于步态高斯图及稀疏表示算法以及基于协同表示算法的识别率。

4 结束语

将 GEI 算法应用到步态识别中可以很好地反映出步态轮廓特征的主要形状，但是该算法没有考虑步态轮廓内边界信息，导致识别率降低。针对此问题，本文提出了使用 Hog 算法与 GEI 算法融合的方法提取步态特征能量图即 SEF-GHEI 能量图，在 SEF-GHEI 能量图的基础上使用协同表示的分类器进行分类识别。为了降低计算复杂度以及得到更有效的步态特征，本文使用 PCA 方法对 SEF-GHEI 步态能量图进行降维，对降维后的 SEF-GHEI 能量图进行训练，最后根据训练得到的投影矩阵 P 计算测试样本的最小重构误差进行分类。实验结果表明，本文算法在单一视角以及跨视角下的三种行走状态的步态识别效果良好，且能达到实时性的要求。由于实际场景更加复杂、监控器拍摄角度以及物体遮掩的原因，对识别造成一定的影响。因此，今后还要增加复杂场景以及跨视角进行实验，以至于更接近现实场景。

参考文献：

[1] Ding M, Fan G. Articulated and generalized Gaussian kernel correlation for human pose estimation [J]. IEEE Tran on Image Processing A Publication of the IEEE Signal Processing Society, 2015, 25 (2): 776-789.

[2] Zhang Y, Pan G, Jia K, *et al.* Accelerometer-based gait recognition by sparse representation of signature points with clusters [J]. IEEE Trans on Cybernetics, 2017, 45 (9): 1864-1875.

[3] Fu L, Zhang J, Huang K. ORGM: occlusion relational graphical model for human pose estimation [J]. IEEE Trans on Image Processing, 2016, 26 (2): 927-941.

[4] Johnson S, Everingham M. Combining discriminative appearance and segmentation cues for articulated human pose estimation [C]// Proc of IEEE International Conference on Computer Vision Workshops. 2010: 405-412.

[5] 林国军，解梅. 一种鲁棒协作表示的人脸识别算法 [J]. 计算机应用研究, 2014, 31 (8): 2520-2522.

[6] Yeoh T W, Aguirre H E, Tanaka K. Clothing-invariant gait recognition using convolutional neural network [C]// Proc of IEEE International Symposium on Intelligent Signal Processing and Communication Systems. 2017: 1-5.

[7] Jia N, Li C T, Sanchez V, *et al.* Fast and robust framework for view-invariant gait recognition [C]// Proc of IEEE International Workshop on Biometrics and Forensics. 2017: 1-6.

[8] Zhang Y, Pan G, Jia K, *et al.* Accelerometer-based gait recognition by sparse representation of signature points with clusters. [J]. IEEE Trans on Cybernetics, 2017, 45 (9): 1864-1875.

[9] Yu T, Wang R. Enhancing scene parsing by transferring structures via efficient low-rank graph matching [C]// Proc of ACM SIGSPATIAL

International Conference on Advances in Geographic Information Systems. 2016.

[10] Lee L, Grimson W E L. Gait analysis for recognition and classification [C]// Proc of IEEE International Conference on Automatic Face and Gesture Recognition. 2002: 148-155.

[11] Bobick A F, Johnson A Y. Gait recognition using static, activity-specific parameters [C]// Proc of IEEE Computer Society Conference on Vision and Pattern Recognition. 2001: 423-430.

[12] 李占利, 孙卓, 杨晓强. 基于步态高斯图及稀疏表示的步态识 [J]. 科学技术与工程, 2017, 17 (4): 250-254.

[13] 杨旗, 薛定宇, 崔建江. 基于稀疏表示的步态识别 [J]. 东北大学学报: 自然科学版, 2012, 33 (1): 43-46.

[14] 李占利, 崔磊磊, 刘金瑄. 基于协同表示的步态识别 [J]. 计算机应用研究, 2016, 33 (9): 2878-2880.

[15] Connie T, Goh M K O, Teoh A B J. A Grassmannian approach to address view change problem in gait recognition [J]. IEEE Trans on Cybernetics, 2017, 47 (6): 1-14.

[16] Eum H, Yoon C, Lee H, *et al.* Continuous human action recognition using depth-MHI-HOG and a spotter model [J]. Sensors, 2015, 15 (3): 51-97.

[17] 杨路明, 曾莹, 曾庆冬, 等. 基于特征融合的步态识别算法研究 [J]. 计算机应用研究, 2008, 25 (7): 2216-2218.

[18] 康利攀, 陈方福, 范自柱, 等. 一种新的拓展稀疏人脸识别算法 [J]. 计算机应用研究, 2016, 33 (3): 937-939.

[19] 崔雪红, 刘云, 常伟, 等. 基于 HOG 特征的步态能量图身份识别算法 [J]. 电子测量技术, 2017, 40 (7): 100-104.

[20] Zhang Y, Pan G, Jia K, *et al.* Accelerometer-based gait recognition by sparse representation of signature points with clusters [J]. IEEE Trans on Cybernetics, 2017, 45 (9): 1864-1875.

[21] Zhang L, Yang M. Sparse representation or collaborative representation: Which helps face recognition? [C]// Proc of IEEE International Conference on Computer Vision. 2012: 471-478.